

160
1632

0110/0240

#6

OIPE

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/843,007

DATE: 08/20/2001

TIME: 12:37:48

RECEIVED

FEB 21 2002

Input Set : C:\PAOLA\09843007.txt

Output Set: N:\CRF3\08202001\I843007.raw

TECH CENTER 1600/2900

SEQUENCE LISTING

4 (1) GENERAL INFORMATION:

6 (i) APPLICANT: Kossmann, Jens
 7 Butcher, Volker
 8 Welsh, Thomas

10 (ii) TITLE OF INVENTION: DNA SEQUENCES ENCODING ENZYMES CAPABLE
 11 OF FACILITATING THE SYNTHESIS OF LINEAR
 12 ALPHA-1,4 GLUCANS IN PLANTS, FUNGI AND
 13 MICROORGANISMS

15 (iii) NUMBER OF SEQUENCES: 4

17 (iv) CORRESPONDENCE ADDRESS:

18 (A) ADDRESSEE: James F. Haley, Jr., c/o Fish & Neave
 19 (B) STREET: 1251 Avenue of the Americas
 20 (C) CITY: New York
 21 (D) STATE: New York
 22 (E) COUNTRY: United States of America
 23 (F) ZIP: 10020

25 (v) COMPUTER READABLE FORM:

26 (A) MEDIUM TYPE: Floppy disk
 27 (B) COMPUTER: IBM PC compatible
 28 (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 29 (D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

31 (vi) CURRENT APPLICATION DATA:

C--> 32 (A) APPLICATION NUMBER: US/09/843,007

C--> 33 (B) FILING DATE: 26-Apr-2001

34 (C) CLASSIFICATION:

39 (vii) PRIOR APPLICATION DATA:

37 (A) APPLICATION NUMBER: 08/737,752

38 (B) FILING DATE:

40 (A) APPLICATION NUMBER: DE P 44 47 388.5

41 (B) FILING DATE: 22-DEC-1994

43 (viii) ATTORNEY/AGENT INFORMATION:

44 (A) NAME: James F. Haley, Jr., Esq.

45 (B) REGISTRATION NUMBER: 27,794

C--> 46 (C) REFERENCE/DOCKET NUMBER: GFB-1

48 (ix) TELECOMMUNICATION INFORMATION:

49 (A) TELEPHONE: (212) 596-9000

50 (B) TELEFAX: (212) 596-9090

53 (2) INFORMATION FOR SEQ ID NO: 1:

55 (i) SEQUENCE CHARACTERISTICS:

56 (A) LENGTH: 2914 base pairs

57 (B) TYPE: nucleic acid

58 (C) STRANDEDNESS: double

59 (D) TOPOLOGY: linear

61 (ii) MOLECULE TYPE: DNA (genomic)

63 (iii) HYPOTHETICAL: NO

65 (iv) ANTI-SENSE: NO

ENTERED

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/843,007

DATE: 08/20/2001

TIME: 12:37:48

Input Set : C:\PAOLA\09843007.txt

Output Set: N:\CRF3\08202001\I843007.raw

```

67      (vi) ORIGINAL SOURCE:
68          (A) ORGANISM: Neisseria polysaccharea
70      (vii) IMMEDIATE SOURCE:
71          (A) LIBRARY: genomic library in pBluescriptII SK
72          (B) CLONE: pNB2
74      (ix) FEATURE:
75          (A) NAME/KEY: CDS
76          (B) LOCATION: 957..2867
79      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
81 GAGTTTTCGCG TTCCCGAACC GAACGTGATG CTTGAGCCGA ACACCTGTCC GGCAAGGCGG      60
83 CTGACCGCCC CCTTTTGCCC CATCGACATC GTAACAATCG GTTTGGTGGC AAGCTCTTTC      120
85 GCTTTGAGCG TGGCAGAAAG CAAAGTCAGC ACGTCTCCG CGCTTTGCGG CATCACCGCA      180
87 ATTTTGACAG TGTCCGCGCC GCAGTCCTTC ATCTGTTTCA GACGGCATAAC GATTTCTTCT      240
89 TCGCGCGCGG TCGCGTGAAA CTCATGATG CAGAGCAGGG CGGCGATGCC GTTTTTTTGA      300
91 GCATGCGCCA CGGCGCGCCG GACGCGGGTT TCGCCGGAAA AAAGCTCGAT ATCGATAATG      360
93 TCGGGCAGGC GGCTTTCAAT CAGCGAGTCG AGCAGTTCAA AATAATAATC GTCCGAACAC      420
95 GGGAACGAGC CGCCTTCGCC ATGCCGTCTG AACGTAAACA GCAGCGGCTT GTCGGGCAGC      480
97 GCGTCGCGGA CGGTCTGCGT GTGGCGCAAT ACTTCGCCGA TGCTGCCCGC GCATTCCAAA      540
99 AAATCGGCGC GGAACCTCGAC GATATCGAAG GGCAGGTTT TGATTTGGTC AAGTACGGCG      600
101 GAAAGTACGG CGGCATCGCG GGCGACAAGC GGCACGGCGA TTTTGGTGCG TCCGCTCCG      660
103 ATAACGGTGT TTTTGACGGT CAGGCTGGTG TGCATGGCGG TTGTTGCGGC TGAAAGGAAC      720
105 GGTAAAGACG CAATTATAGC AAAGGCACAG GCAATGTTTC AGACGGCATT TCTGTGCGGC      780
107 CGGCTTGATA TGAATCAAGC AGCATCCGCA TATCGGAATG CAGACTTGGC ACAAGCCCTG      840
109 TCTTTTCTAG TCAGTCCGCA GTTCTTGAGC TATGATTGCA CGACACGCCC TACACGGCAT      900
111 TTGCAGGATA CGGCGGCAGA CCGCCGGTCG GAAACTTCAG AATCGGAGCA GGCATC      956
113 ATG TTG ACC CCC ACG CAG CAA GTC GGT TTG ATT TTA CAG TAC CTC AAA      1004
114 Met Leu Thr Pro Thr Gln Gln Val Gly Leu Ile Leu Gln Tyr Leu Lys
115 1 5 10 15
117 ACA CGC ATC TTG GAC ATC TAC ACG CCC GAA CAG CGC GCC GGC ATC GAA      1052
118 Thr Arg Ile Leu Asp Ile Tyr Thr Pro Glu Gln Arg Ala Gly Ile Glu
119 20 25 30
121 AAA TCC GAA GAC TGG CGG CAG TTT TCG CGC CGC ATG GAT ACG CAT TTC      1100
122 Lys Ser Glu Asp Trp Arg Gln Phe Ser Arg Arg Met Asp Thr His Phe
123 35 40 45
125 CCC AAA CTG ATG AAC GAA CTC GAC AGC GTG TAC GGC AAC AAC GAA GCC      1148
126 Pro Lys Leu Met Asn Glu Leu Asp Ser Val Tyr Gly Asn Asn Glu Ala
127 50 55 60
129 CTG CTG CCT ATG CTG GAA ATG CTG CTG GCG CAG GCA TGG CAA AGC TAT      1196
130 Leu Leu Pro Met Leu Glu Met Leu Leu Ala Gln Ala Trp Gln Ser Tyr
131 65 70 75 80
133 TCC CAA CGC AAC TCA TCC TTA AAA GAT ATC GAT ATC GCG CGC GAA AAC      1244
134 Ser Gln Arg Asn Ser Ser Leu Lys Asp Ile Asp Ile Ala Arg Glu Asn
135 85 90 95
137 AAC CCC GAT TGG ATT TTG TCC AAC AAA CAA GTC GGC GGC GTG TGC TAC      1292
138 Asn Pro Asp Trp Ile Leu Ser Asn Lys Gln Val Gly Gly Val Cys Tyr
139 100 105 110
141 GTT GAT TTG TTT GCC GGC GAT TTG AAG GGC TTG AAA GAT AAA ATT CCT      1340
142 Val Asp Leu Phe Ala Gly Asp Leu Lys Gly Leu Lys Asp Lys Ile Pro
143 115 120 125

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/843,007

DATE: 08/20/2001

TIME: 12:37:48

Input Set : C:\PAOLA\09843007.txt

Output Set: N:\CRF3\08202001\I843007.raw

145	TAT	TTT	CAA	GAG	CTT	GGT	TTG	ACT	TAT	CTG	CAC	CTG	ATG	CCG	CTG	TTT	1388
146	Tyr	Phe	Gln	Glu	Leu	Gly	Leu	Thr	Tyr	Leu	His	Leu	Met	Pro	Leu	Phe	
147		130					135					140					
149	AAA	TGC	CCT	GAA	GGC	AAA	AGC	GAC	GGC	GGC	TAT	GCG	GTC	AGC	AGC	TAC	1436
150	Lys	Cys	Pro	Glu	Gly	Lys	Ser	Asp	Gly	Gly	Tyr	Ala	Val	Ser	Ser	Tyr	
151	145					150					155					160	
153	CGC	GAT	GTC	AAT	CCG	GCA	CTG	GGC	ACA	ATA	GGC	GAC	TTG	CGC	GAA	GTC	1484
154	Arg	Asp	Val	Asn	Pro	Ala	Leu	Gly	Thr	Ile	Gly	Asp	Leu	Arg	Glu	Val	
155					165					170					175		
157	ATT	GCT	GCG	CTG	CAC	GAA	GCC	GGC	ATT	TCC	GCC	GTC	GTC	GAT	TTT	ATC	1532
158	Ile	Ala	Ala	Leu	His	Glu	Ala	Gly	Ile	Ser	Ala	Val	Val	Asp	Phe	Ile	
159			180						185					190			
161	TTC	AAC	CAC	ACC	TCC	AAC	GAA	CAC	GAA	TGG	GCG	CAA	CGC	TGC	GCC	GCC	1580
162	Phe	Asn	His	Thr	Ser	Asn	Glu	His	Glu	Trp	Ala	Gln	Arg	Cys	Ala	Ala	
163			195					200					205				
165	GGC	GAC	CCG	CTT	TTC	GAC	AAT	TTC	TAC	TAT	ATT	TTC	CCC	GAC	CGC	CGG	1628
166	Gly	Asp	Pro	Leu	Phe	Asp	Asn	Phe	Tyr	Tyr	Ile	Phe	Pro	Asp	Arg	Arg	
167		210					215					220					
169	ATG	CCC	GAC	CAA	TAC	GAC	CGC	ACC	CTG	CGC	GAA	ATC	TTC	CCC	GAC	CAG	1676
170	Met	Pro	Asp	Gln	Tyr	Asp	Arg	Thr	Leu	Arg	Glu	Ile	Phe	Pro	Asp	Gln	
171	225					230					235					240	
173	CAC	CCG	GGC	GGC	TTC	TCG	CAA	CTG	GAA	GAC	GGA	CGC	TGG	GTG	TGG	ACG	1724
174	His	Pro	Gly	Gly	Phe	Ser	Gln	Leu	Glu	Asp	Gly	Arg	Trp	Val	Trp	Thr	
175					245					250					255		
177	ACC	TTC	AAT	TCC	TTC	CAA	TGG	GAC	TTG	AAT	TAC	AGC	AAC	CCG	TGG	GTA	1772
178	Thr	Phe	Asn	Ser	Phe	Gln	Trp	Asp	Leu	Asn	Tyr	Ser	Asn	Pro	Trp	Val	
179			260					265						270			
181	TTC	CGC	GCA	ATG	GCG	GGC	GAA	ATG	CTG	TTC	CTT	GCC	AAC	TTG	GGC	GTT	1820
182	Phe	Arg	Ala	Met	Ala	Gly	Glu	Met	Leu	Phe	Leu	Ala	Asn	Leu	Gly	Val	
183			275					280					285				
185	GAC	ATC	CTG	CGT	ATG	GAT	GCG	GTT	GCC	TTT	ATT	TGG	AAA	CAA	ATG	GGG	1868
186	Asp	Ile	Leu	Arg	Met	Asp	Ala	Val	Ala	Phe	Ile	Trp	Lys	Gln	Met	Gly	
187		290					295					300					
189	ACA	AGC	TGC	GAA	AAC	CTG	CCG	CAG	GCG	CAC	GCC	CTC	ATC	CGC	GCG	TTC	1916
190	Thr	Ser	Cys	Glu	Asn	Leu	Pro	Gln	Ala	His	Ala	Leu	Ile	Arg	Ala	Phe	
191	305					310					315					320	
193	AAT	GCC	GTT	ATG	CGT	ATT	GCC	GCG	CCC	GCC	GTG	TTC	TTC	AAA	TCC	GAA	1964
194	Asn	Ala	Val	Met	Arg	Ile	Ala	Ala	Pro	Ala	Val	Phe	Phe	Lys	Ser	Glu	
195					325					330					335		
197	GCC	ATC	GTC	CAC	CCC	GAC	CAA	GTC	GTC	CAA	TAC	ATC	GGG	CAG	GAC	GAA	2012
198	Ala	Ile	Val	His	Pro	Asp	Gln	Val	Val	Gln	Tyr	Ile	Gly	Gln	Asp	Glu	
199			340					345						350			
201	TGC	CAA	ATC	GGT	TAC	AAC	CCC	CTG	CAA	ATG	GCA	TTG	TTG	TGG	AAC	ACC	2060
202	Cys	Gln	Ile	Gly	Tyr	Asn	Pro	Leu	Gln	Met	Ala	Leu	Leu	Trp	Asn	Thr	
203			355					360					365				
205	CTT	GCC	ACG	CGC	GAA	GTC	AAC	CTG	CTC	CAT	CAG	GCG	CTG	ACC	TAC	CGC	2108
206	Leu	Ala	Thr	Arg	Glu	Val	Asn	Leu	Leu	His	Gln	Ala	Leu	Thr	Tyr	Arg	
207		370					375					380					
209	CAC	AAC	CTG	CCC	GAG	CAT	ACC	GCC	TGG	GTC	AAC	TAC	GTC	CGC	AGC	CAC	2156

RAW SEQUENCE LISTING

DATE: 08/20/2001

PATENT APPLICATION: US/09/843,007

TIME: 12:37:48

Input Set : C:\PAOLA\09843007.txt

Output Set: N:\CRF3\08202001\I843007.raw

210	His	Asn	Leu	Pro	Glu	His	Thr	Ala	Trp	Val	Asn	Tyr	Val	Arg	Ser	His	
211	385					390					395					400	
213	GAC	GAC	ATC	GGC	TGG	ACG	TTT	GCC	GAT	GAA	GAC	GCG	GCA	TAT	CTG	GGC	2204
214	Asp	Asp	Ile	Gly	Trp	Thr	Phe	Ala	Asp	Glu	Asp	Ala	Ala	Tyr	Leu	Gly	
215				405					410						415		
217	ATA	AGC	GGC	TAC	GAC	CAC	CGC	CAA	TTC	CTC	AAC	CGC	TTC	TTC	GTC	AAC	2252
218	Ile	Ser	Gly	Tyr	Asp	His	Arg	Gln	Phe	Leu	Asn	Arg	Phe	Phe	Val	Asn	
219				420					425					430			
221	CGT	TTC	GAC	GGC	AGC	TTC	GCT	CGT	GGC	GTA	CCG	TTC	CAA	TAC	AAC	CCA	2300
222	Arg	Phe	Asp	Gly	Ser	Phe	Ala	Arg	Gly	Val	Pro	Phe	Gln	Tyr	Asn	Pro	
223				435					440					445			
225	AGC	ACA	GGC	GAC	TGC	CGT	GTC	AGT	GGT	ACA	GCC	GCG	GCA	TTG	GTC	GGC	2348
226	Ser	Thr	Gly	Asp	Cys	Arg	Val	Ser	Gly	Thr	Ala	Ala	Ala	Leu	Val	Gly	
227		450					455						460				
229	TTG	GCG	CAA	GAC	GAT	CCC	CAC	GCC	GTT	GAC	CGC	ATC	AAA	CTC	TTG	TAC	2396
230	Leu	Ala	Gln	Asp	Asp	Pro	His	Ala	Val	Asp	Arg	Ile	Lys	Leu	Leu	Tyr	
231	465					470					475					480	
233	AGC	ATT	GCT	TTG	AGT	ACC	GGC	GGT	CTG	CCG	CTG	ATT	TAC	CTA	GGC	GAC	2444
234	Ser	Ile	Ala	Leu	Ser	Thr	Gly	Gly	Leu	Pro	Leu	Ile	Tyr	Leu	Gly	Asp	
235				485					490						495		
237	GAA	GTG	GGT	ACG	CTC	AAT	GAC	GAC	GAC	TGG	TCG	CAA	GAC	AGC	AAT	AAG	2492
238	Glu	Val	Gly	Thr	Leu	Asn	Asp	Asp	Asp	Trp	Ser	Gln	Asp	Ser	Asn	Lys	
239				500					505					510			
241	AGC	GAC	GAC	AGC	CGT	TGG	GCG	CAC	CGT	CCG	CGC	TAC	AAC	GAA	GCC	CTG	2540
242	Ser	Asp	Asp	Ser	Arg	Trp	Ala	His	Arg	Pro	Arg	Tyr	Asn	Glu	Ala	Leu	
243				515					520					525			
245	TAC	GCG	CAA	CGC	AAC	GAT	CCG	TCG	ACC	GCA	GCC	GGG	CAA	ATC	TAT	CAG	2588
246	Tyr	Ala	Gln	Arg	Asn	Asp	Pro	Ser	Thr	Ala	Ala	Gly	Gln	Ile	Tyr	Gln	
247		530					535						540				
249	GGC	TTG	CGC	CAT	ATG	ATT	GCC	GTC	CGC	CAA	AGC	AAT	CCG	CGC	TTC	GAC	2636
250	Gly	Leu	Arg	His	Met	Ile	Ala	Val	Arg	Gln	Ser	Asn	Pro	Arg	Phe	Asp	
251	545					550					555				560		
253	GGC	GGC	AGG	CTG	GTT	ACA	TTC	AAC	ACC	AAC	AAC	AAG	CAC	ATC	ATC	GGC	2684
254	Gly	Gly	Arg	Leu	Val	Thr	Phe	Asn	Thr	Asn	Asn	Lys	His	Ile	Ile	Gly	
255				565					570					575			
257	TAC	ATC	CGC	AAC	AAT	GCG	CTT	TTG	GCA	TTC	GGT	AAC	TTC	AGC	GAA	TAT	2732
258	Tyr	Ile	Arg	Asn	Asn	Ala	Leu	Leu	Ala	Phe	Gly	Asn	Phe	Ser	Glu	Tyr	
259				580					585					590			
261	CCG	CAA	ACC	GTT	ACC	GCG	CAT	ACC	CTG	CAA	GCC	ATG	CCC	TTC	AAG	GCG	2780
262	Pro	Gln	Thr	Val	Thr	Ala	His	Thr	Leu	Gln	Ala	Met	Pro	Phe	Lys	Ala	
263				595					600					605			
265	CAC	GAC	CTC	ATC	GGT	GGC	AAA	ACT	GTC	AGC	CTG	AAT	CAG	GAT	TTG	ACG	2828
266	His	Asp	Leu	Ile	Gly	Gly	Lys	Thr	Val	Ser	Leu	Asn	Gln	Asp	Leu	Thr	
267		610					615						620				
269	CTT	CAG	CCC	TAT	CAG	GTC	ATG	TGG	CTC	GAA	ATC	GCC	TGA	CGC	ACG	CTT	2877
270	Leu	Gln	Pro	Tyr	Gln	Val	Met	Trp	Leu	Glu	Ile	Ala	*				
271	625					630							635				
273	CCAAATGCCG	TCTGAACCGT	TTCAGACGGC	ATTGCG													2914
276	(2)	INFORMATION	FOR	SEQ	ID	NO:	2:										

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/843,007

DATE: 08/20/2001

TIME: 12:37:48

Input Set : C:\PAOLA\09843007.txt

Output Set: N:\CRF3\08202001\I843007.raw

```

278      (i) SEQUENCE CHARACTERISTICS:
279          (A) LENGTH: 636 amino acids
280          (B) TYPE: amino acid
281          (D) TOPOLOGY: linear
283      (ii) MOLECULE TYPE: protein
285      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
287 Met Leu Thr Pro Thr Gln Gln Val Gly Leu Ile Leu Gln Tyr Leu Lys
288   1          5          10          15
290 Thr Arg Ile Leu Asp Ile Tyr Thr Pro Glu Gln Arg Ala Gly Ile Glu
291          20          25          30
293 Lys Ser Glu Asp Trp Arg Gln Phe Ser Arg Arg Met Asp Thr His Phe
294          35          40          45
296 Pro Lys Leu Met Asn Glu Leu Asp Ser Val Tyr Gly Asn Asn Glu Ala
297   50          55          60
299 Leu Leu Pro Met Leu Glu Met Leu Leu Ala Gln Ala Trp Gln Ser Tyr
300  65          70          75          80
302 Ser Gln Arg Asn Ser Ser Leu Lys Asp Ile Asp Ile Ala Arg Glu Asn
303          85          90          95
305 Asn Pro Asp Trp Ile Leu Ser Asn Lys Gln Val Gly Gly Val Cys Tyr
306          100         105         110
308 Val Asp Leu Phe Ala Gly Asp Leu Lys Gly Leu Lys Asp Lys Ile Pro
309          115         120         125
311 Tyr Phe Gln Glu Leu Gly Leu Thr Tyr Leu His Leu Met Pro Leu Phe
312          130         135         140
314 Lys Cys Pro Glu Gly Lys Ser Asp Gly Gly Tyr Ala Val Ser Ser Tyr
315 145         150         155         160
317 Arg Asp Val Asn Pro Ala Leu Gly Thr Ile Gly Asp Leu Arg Glu Val
318          165         170         175
320 Ile Ala Ala Leu His Glu Ala Gly Ile Ser Ala Val Val Asp Phe Ile
321          180         185         190
323 Phe Asn His Thr Ser Asn Glu His Glu Trp Ala Gln Arg Cys Ala Ala
324          195         200         205
326 Gly Asp Pro Leu Phe Asp Asn Phe Tyr Tyr Ile Phe Pro Asp Arg Arg
327          210         215         220
329 Met Pro Asp Gln Tyr Asp Arg Thr Leu Arg Glu Ile Phe Pro Asp Gln
330 225         230         235         240
332 His Pro Gly Gly Phe Ser Gln Leu Glu Asp Gly Arg Trp Val Trp Thr
333          245         250         255
335 Thr Phe Asn Ser Phe Gln Trp Asp Leu Asn Tyr Ser Asn Pro Trp Val
336          260         265         270
338 Phe Arg Ala Met Ala Gly Glu Met Leu Phe Leu Ala Asn Leu Gly Val
339          275         280         285
341 Asp Ile Leu Arg Met Asp Ala Val Ala Phe Ile Trp Lys Gln Met Gly
342          290         295         300
344 Thr Ser Cys Glu Asn Leu Pro Gln Ala His Ala Leu Ile Arg Ala Phe
345 305         310         315         320
347 Asn Ala Val Met Arg Ile Ala Ala Pro Ala Val Phe Phe Lys Ser Glu
348          325         330         335
350 Ala Ile Val His Pro Asp Gln Val Val Gln Tyr Ile Gly Gln Asp Glu

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/843,007

DATE: 08/20/2001

TIME: 12:37:49

Input Set : C:\PAOLA\09843007.txt

Output Set: N:\CRF3\08202001\I843007.raw

L:32 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:33 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]
L:46 M:220 C: Keyword misspelled or invalid format, [(C) REFERENCE/DOCKET NUMBER:]